Democracy and the Logic of Political Survival [§]

Kevin A. Clarke [†] and Randall W. Stone [‡]
University of Rochester

Abstract

Although democracy is a key concept in political science, debate continues over definitions and mechanisms. Bueno de Mesquita *et al.* (2003) make the important claim that most of democracy's effects are in fact due to something conceptually simpler and empirically easier than democracy: the size of the minimum winning coalition that selects the leader. The argument is intuitively appealing and supported by extensive data analysis. Unfortunately, the statistical technique they use induces omitted variable bias into their results. They argue that they need to control for democracy, but their estimation procedure treats democracy as an omitted variable. When we reestimate their regressions controlling for democracy, most of their important findings do not survive. As a final check, we replicate their analysis using more appropriate measures derived by Powell (2000). The results again fail to support the predictions of the selectorate model.

October 31, 2006

[§]An earlier version of this paper was given at the Watson Center for Conflict and Cooperation at the University of Rochester, October 2006; we thank the participants for their comments. We also thank Hein Goemans, Gretchen Helmke, Elena McLean, Martin Steinwand, and especially Bing Powell and John Jackson for helpful comments and advice. We also thank Bruce Bueno de Mesquita, Alastair Smith, Randy Siverson, and Jim Morrow for sharing their data and code. All errors remain our own.

[†]Assistant Professor, Department of Political Science, University of Rochester, Rochester, NY 14627-0146. Tel.: (585) 275-5217; Fax: (585) 271-1616; Email: `kevin.clarke@rochester.edu`.

[‡]Corresponding author. Associate Professor, Department of Political Science, University of Rochester, Rochester, NY 14627-0146. Tel.: (585) 273-4761; Fax: (585) 271-1616; Email: `randall.stone@rochester.edu`.

1 Introduction

Democracy is an important concept in both comparative and international relations research. Scholars have linked it to patterns of war and peace, alliance formation, conflict escalation, and even success in war. Democracy is apparently associated with economic development, successful restructuring and reform, reduction of inequality, and improvements in a wide array of social indicators. It has been linked to political stability, lower rates of military coups, and the non-violent resolution of domestic conflicts. While each of these claims is subject to challenge, the existence of a substantial literature on each of them underlines the importance of the questions. What, however, is the mechanism by which democracy is supposed to generate all of these benefits? Is democracy a matter of norms, behavior, or institutions? Although democracy is a widely used concept, there is no consensus on its theoretical underpinnings.

It is for these reasons that the appearance of *The Logic of Political Survival* (Bueno de Mesquita, Smith, Siverson, and Morrow 2003) was hailed as a breakthrough.[1] The book provides a simple explanation for the importance of democracy as a theoretical concept: democracies provide different incentives to leaders because democracies have different reselection institutions. All leaders want to survive in office, and the ability to do so depends upon the size of the selectorate (S, the set of politically relevant principals who choose the leader) and the size of the minimum winning coalition of selectors (W) determined by a particular constellation of institutions. Leaders raise resources through taxation and spend them to provide public goods, private benefits to their supporters, and luxuries for themselves. Reselection mechanisms that privilege the median member of the population create incentives to provide public goods; mechanisms that enfranchise only a narrow clique instead create incentives to provide private benefits to those who have the real power. Further, the amount of resources that the leader is able to appropriate for herself—and conversely, the incentive to overthrow her—depends upon the ratio between the size of the winning coalition and the size of the selectorate (W/S) because this determines how indispensable is each member of the coalition.

[1] A version of the formal model originally appeared in Bueno de Mesquita, Smith, Siverson, and Morrow 1999.

The authors argue that the selectorate mechanism is analytically distinct from the concept of democracy, but captures to a significant degree the features of democracy that make it an appealing and effective form of government. Bueno de Mesquita *et al.* advance this claim by testing the predictions of the selectorate model against a wide array of dependent variables. They derive their measures of W and W/S from commonly used measures of democracy, notably the Polity IV score, so W and W/S are highly correlated with the polity score. The authors are sensitive, therefore, to the possible criticism that W and W/S are no more than proxies for democracy, and take pains to show that their effects are robust to models that include polity scores as controls. In fact, one of their key findings is that W and W/S often outperform democracy when included in equations containing democracy. Concluding Chapter 5, they write:

> Characteristics of democracy that are independent of W and S prove to have much less consistent relationships with the dependent variables assessed here than does coalition size [W] or the loyalty norm [W/S]. Neither democracy nor income effects are sufficient to explain the specific predicted impact that coalition size and the loyalty norm have on policy choices (Bueno de Mesquita *et al.* 2003, 214).

If correct, this claim has important implications for the long-standing debate about whether democracy should be understood primarily in terms of attitudes and behavior (e.g. Almond and Verba 1963, Dahl 1971, Putnam 1994, Przeworski, Alvarez, Cheibub, and Limongi 2000) or in terms of institutions (e.g. Shugart and Carey 1992, Cox 1997, Lijphart 1999, Powell 2000). If attitudes, culture, ideology or equilibrium expectations have independent causal weight, it is unlikely that particular constellations of institutions will adequately account for the variation in outcomes that we broadly associate with democratic politics. If, on the other hand, the differences between democratic and authoritarian politics and the differences in politics among democracies are reducible to institutional incentives, then institutions and not behaviors are the crux of the matter. The difficulty that the institutional theory of democracy has faced has been the bewildering diversity of democratic institutions; Lijphart (1999), for example, isolates ten dimensions of variation, and admits that this is a simplification. If Bueno de Mesquita and his coauthors are able to reduce the dimensionality of the problem to one or two

simple concepts, and thereby capture most of the variation in democratic outcomes, this is a major achievement.

We demonstrate that this conclusion is incorrect. In fact, the size of a polity's selectorate and winning coalition are often not associated with the decisions made by leaders or are associated in ways that are opposed to the model's predictions. Bueno de Mesquita *et al.* are misled due to their use of an estimation that induces omitted variable bias: instead of controlling for democracy in their regressions, the authors control for a variable called *democratic residuals*, which are the residuals from an auxiliary regression of democracy on W and S. The effect of this technique is to exaggerate the importance of W or W/S by increasing its coefficient and decreasing its standard error. As we prove, the induced bias is equivalent to omitted variable bias (treating, in this case, democracy as the omitted variable), and the large sample sizes in the study do not decrease the bias as the estimator is also inconsistent.

The correct way to test the predictions of the selectorate theory is to control for democracy, not *democratic residuals.* Therefore, we reestimate the main tables from Chapters 4 and 5 while controlling for democracy.[2] Their key findings do not survive. We conclude that W and W/S owe most of their apparent effect to their correlation with democracy, and their strong effects in models that control for democratic residuals are artifacts of the biased estimator.

It is important to determine whether the results we report could be attributed to poor data. Bueno de Mesquita *et al.* are quick to emphasize the "crude and primitive" nature of their measures, which do not measure the quantities of interest. Their measures of W and S are derived from components of the Polity score combined with a dummy variable derived from a variable developed by Arthur Banks. They argue that their approximations "should, however, be adequate to evaluate whether the central tendencies of politics are aligned with the expectations that follow from the selectorate model" [133]. Since their measures do not measure coalition or selectorate size directly, however, this need not be the case. Consequently, we reestimate their results again using a theoretically more appropriate measure of their key

[2]We would like to emphasize that Bruce Bueno de Mesquita and his coauthors deserve praise as models of good scholarly practice for making all of their data and statistical routines (Stata .do files) publicly available on the internet (`http://www.nyu.edu/gsas/dept/politics/data/bdm2s2/Logic.htm`). It would have been difficult, if not impossible, to replicate their results without their assistance.

independent variable. Again, we find that the results do not support the expectations of their model.

The paper is organized as follows. In Section 2, we discuss Bueno de Mesquita *et al.*'s approach and provide an example drawn from the book, and in Section 3, we demonstrate analytically that their estimator is biased and inconsistent. In Section 4, we replicate their analyses and rerun them using democracy, not democratic residuals, as a control. In Section 5, we replicate their analyses again, using a theoretically more appropriate measure of the key independent variable. The final section concludes.

2 The Logic of Political Survival

Bueno de Mesquita *et al.* (2003) employ a wide range of dependent variables in testing their model, but to give a flavor for their results and what is incorrect about their analysis, we initially focus on a representative example—public expenditures. The argument is as follows. Leaders are motivated by the desire to survive in office, and they assure reselection by providing private and public goods. Private goods are the instrument of choice in political systems with small winning coalitions because private goods disproportionately enrich the disproportionately influential. Public goods, on the other hand, are preferred in systems that require that leaders win election because public goods provide the highest average benefit at the lowest average cost. Leaders earn more rents when their minimum winning coalitions are small because it is cheaper to enrich a few than to provide general welfare for many. Leaders earn even more rents if their small group of followers can be recruited from a large pool because that keeps their supporters insecure and therefore less demanding. One of the most direct implications of their model, therefore, is that institutions that call for large winning coalitions should be associated with more effort to produce public goods. Holding coalition size constant, larger selectorates should be associated with fewer public goods.

One of the key tests of the model, therefore, is based upon the level of public expenditures. They write:

> According to the selectorate theory, spending decisions are strategic. While leaders want to retain as many resources as possible

> for their own discretionary use, institutional arrangements influence their ability to do so without jeopardizing their tenure in office. Because the loyalty norm created by the focus on private goods and the risk of exclusion from future coalitions is strong in small-coalition systems, leaders can provide low levels of rewards and still adequately match the best possible offer of a challenger. Leaders in such systems need not spend much to survive.
>
> This is not true for those who head large-coalition governments with their concomitantly weak loyalty norm. Leaders in such systems need to work hard to provide benefits for their supporters and cannot afford to skim off resources if they wish to survive. Ignoring for the moment the possibility of deficit spending, these incentives mean that when the coalition is small, government expenditures are expected to be relatively small, and when the coalition is larger, so are government expenditures. Small-coalition leaders can survive spending a smaller quantity of what they collect than can large-coalition leaders (Bueno de Mesquita *et al.* 2003, 161-162).

The authors test this proposition by regressing expenditures as a proportion of GDP solely on W or W/S, and the results support their expectations. The authors recognize, of course, that this bare bones specification is incomplete and suffers from omitted variable bias. Democracies, for example, might spend more than autocracies for reasons other than the logic of the selectorate model. At the same time, governments in wealthier countries can spend a greater percentage of their GDP. Thus, Bueno de Mesqauta *et al.* run additional regressions with these variables as controls, and it is here that their analysis goes off track. Instead of controlling for democracy and GDP, they attempt to control for these effects by using the residuals from two other regressions. The first auxiliary equation regresses democracy on W and S, and the second regresses the natural logarithm of per capita income on W and S.[3]

The technique used by Bueno de Mesquita *et al.*, substituting residuals for

[3]In the text, Bueno de Mesquita *et al.* write, "To assess the independent impact of democracy on our dependent variables, we create a variable called *WS:DemRes* by regressing W and S on *Democracy* and then saving the residuals...." [137]. Their .do files, however, show that they actually regress *Democracy* on W and S.

Table 1: Expenditures, $n = 2373$

	Residuals		Controls	
Variable	Coef.	s.e.	Coef.	s.e.
W	2.32	(0.99)	-4.47	(1.74)
S	3.10	(0.85)	4.35	(0.87)
ln(Population)	-1.90	(0.13)	-1.90	(0.13)
Democracy residuals	2.83	(1.15)		
ln(GDP) residuals	1.07	(0.20)		
Democracy			2.83	(1.15)
ln(GDP)			1.07	(0.20)
Constant	53.89	(2.19)	47.49	(2.69)

control variables, serves to exaggerate the size of the coefficients on W and W/S because both democracy and the log of income are positively correlated with W. (We explain why this correlation matters in the next section.) Table 1 displays our replication of Bueno de Mesquita *et al.*'s government expenditure results using their data and batch files. Table 1 also displays our reestimation of these models correctly controlling for *democracy* and the *log of GDP*. The results that appeared to have supported the selectorate theory evaporate.

The first column of Table 1 replicates Model 2 in Table 4.6 of Bueno de Mesquita *et al.* (2003, 163). The coefficient for W is in the expected direction—larger coalitions exhibit higher levels of expenditures—and it is statistically significant ($p = 0.02$). "Societies with high per capita incomes spend more," they write. "However, even after controlling for these effects (or the independent, positive, and significant impact of democracy), large coalition size and a weak loyalty norm are still powerful indicators of higher government spending" [163]. In the second column, we replace the residuals of democracy and income with the original variables. Not only do the results no longer support the selectorate model; they actually move in the opposite direction. The coefficient on W reverses signs—larger coalitions are associated with lower levels of expenditures—and it is statistically significant ($p = 0.01$).[4] The evidence that Bueno de Mesquita *et al.* point to is an artifact of their biased estimation procedure.

[4]The coefficients on *democracy* and *ln(GDP)* do not change when we replace the residuals with the actual variables. We explain why this is the case in the next section.

Bueno de Mesquita *et al.* construe their results to indicate that the variables that they have constructed, W and W/S, have important effects that are independent of democracy and the level of economic development. They write:

> We provided a set of demanding models that include region-year fixed effects and that isolate the effects of W and S, as well as other remaining, independent effects of democracy or autocracy, and of income. ...Despite the demands of these tests and the crudity of our indicators for coalition size and selectorate size, the evidence consistently provided encouragement for the view that the selectorate theory offers a plausible account of the ties between governance structure and key aspects of economic and political performance (Bueno de Mesquita *et al.* 2003, 171).

This conclusion is false. The apparently robust performance of W and W/S is due to the biased estimation procedure. To the extent that W or W/S explain the dependent variables studied, it is generally because they are correlated with democracy and the log of GDP. When these two variables are included in the specification, the effects of W and W/S disappear, reverse signs, or are substantially attenuated.

There are many more results reported in Chapters 4 and 5 of Bueno de Mesquita *et al.* (2003) that are reasonable tests of the selectorate model, and we review these systematically in Section 4, below. When estimated correctly, the results of these tests do not provide broad support for the conclusion that the variables W and W/S have any impact independent of democracy, income, and several other control variables that Bueno de Mesquita *et al.* use. These data provide little evidence to support the selectorate model.

3 Why the Estimator is Biased

The technique used by Bueno de Mesquita *et al.* induces omitted variable bias. Known as "residualization," the procedure is well-known in certain subfields of political science, particularly American politics, although recent uses, such as Hicken *et al.* (2005), cite Bueno de Mesquita *et al.* (2003)

directly. In the discussion that follows, we demonstrate how the residualization technique induces omitted variable bias, discuss how and why Bueno de Mesquita *et al.* are mistaken in their use of residualization, and briefly document other uses of the procedure.

Intuitively, the problem with Bueno de Mesquita *et al.*'s approach is that the shared variation between, for example, W, *democracy*, and the dependent variable is added to the shared variation between W and the dependent variable. The effect is to bias the coefficient on W and lower the standard error (due to the increased information). To be specific, consider the regression of expenditures, E, on W and democracy, D. This regression is depicted in Figure 1 using Kennedy's Ballentine (Kennedy 1982, 63).[5] There are three areas of interest: area a, which is the shared variation between E and W, area b, which is the shared variation between E and D, and area c, which is the shared variation between E, D, and W. The shared variation between E and W, area a, is used to estimate the coefficient and standard error on winning coalition size. The shared variation between E and D, area b, is used to estimate the coefficient and standard error on democracy. The procedure calls for regressing D on W in an auxiliary regression and then replacing D in the main regression with the residuals from the auxiliary regression. These residuals are circle D in Figure 1 minus area $c + d$. As W and these residuals are uncorrelated by construction, the independent variables in the main equation are orthogonal to one another. What then happens to area c, the shared variation between E, D, and W? It is added to area a, the shared variation between E and W. There are three effects.

- The coefficient on W is biased because the regression has been told that there is greater shared variation between E and W ($a + c$ as opposed to just a).
- This increased information leads to an attenuation of the standard error for the coefficient on W.
- As the shared variation between E and D, area b, remains unchanged by replacing D with the residuals, the coefficient on the residuals is the same as if D had been included in the regression; it is unbiased. (See, for example, the coefficients on *democracy* and *ln(GDP)* in Table 1.)

[5] We also draw on Kennedy's discussion in the remainder of this paragraph.

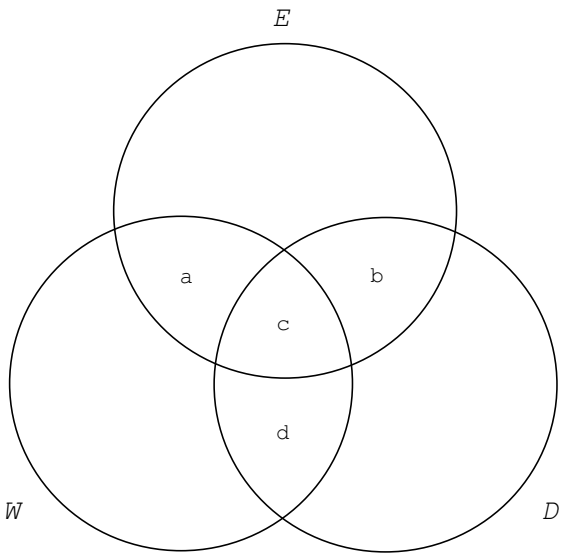

Figure 1: Ballentine (Kennedy 1982, 63)

The true danger of this technique is seen when we realize that the effect on W's estimated coefficient *is precisely the same as if we had omitted D from the regression altogether.* The bias on W's estimated coefficient is omitted variable bias. For the sake of notational simplicity, let us assume a model with only two independent variables: winning coalition size, $\mathbf{w}$, and democracy, $\mathbf{d}$,

$$\mathbf{y} = \beta_0 + \mathbf{w}\beta_1 + \mathbf{d}\beta_2 + \boldsymbol{\epsilon},$$

where $\mathbf{y}$, $\mathbf{w}$, $\mathbf{d}$, and $\boldsymbol{\epsilon}$ are vectors of length n. As Bueno de Mesquita *et al.* worry about the correlation between winning coalition size and democracy, they replace democracy in the equation above with the residuals from a regression of democracy on winning coalition size,

$$\mathbf{y} = \delta_0 + \mathbf{w}\delta_1 + \hat{\boldsymbol{\nu}}\beta_2 + \boldsymbol{\epsilon}^6,$$

[6]We have left the coefficient on $\hat{\boldsymbol{\nu}}$ as β_2 to emphasize that the residualization technique does not change this coefficient.

where $\hat{\boldsymbol{\nu}}$ are the residuals from the auxiliary regression,

$$\mathbf{d} = \gamma_0 + \mathbf{w}\gamma_1 + \boldsymbol{\nu}.$$

The expected value of this estimator, under the usual assumption that the expectation of the error term is zero, $E[\boldsymbol{\epsilon}] = \mathbf{0}$, is

$$E[\hat{\delta}_1] = \delta_1 + (\mathbf{w}'\mathbf{w})^{-1}\mathbf{w}'\mathbf{d}\beta_2,$$

The bias for the estimated coefficient on W is therefore

$$B(\hat{\delta}_1; \delta_1) = E[\hat{\delta}_1] - \delta_1 = (\mathbf{w}'\mathbf{w})^{-1}\mathbf{w}'\mathbf{d}\beta_2. \tag{1}$$

Equation 1 is precisely the bias that would result if democracy were left out of the specification completely (Greene 2003, 148-49). The procedure amounts to generating omitted variable bias on the coefficient of interest while still getting an unbiased estimate on the omitted variable. The formal derivation of equivalence between this estimator and omitted variable bias, as well as additional results concerning the variance and mean squared error comparisons, are in Appendix A.

This equivalence cannot be observed directly in the results Bueno de Mesquita *et al.* report. When they move from basic models that include only W or W/S to models that include the residuals, additional covariates, such as *log(population)* and *log(GDP)* are included as well. We can see the equivalence if we compare a regression of *expenditures* on W and the democracy residuals to a regression of *expenditures* on just W. These results are in Table 2, and we can see that the estimated coefficients on W are precisely the same in both cases.

Table 2: Expenditures, $n = 2500$

	Residuals		Controls	
Variable	Coef.	s.e.	Coef.	s.e.
W	7.76	(0.954)	7.76	(0.959)
Democracy residuals	-9.03	(1.636)		
Constant	23.43	(0.683)	23.43	(0.687)

Democracy, it turns out, is not the only variable for which Bueno de Mesquita *et al.* substitute residuals. They use, in various specifications, the "residual impact" of GDP, civil liberties, political rights, investment, and savings. How these biases interact with one another is demonstrated easily with an example. Consider the results in Table 3, in which we replicate and reestimate the results of a regression of savings on W and controls (the replication is of model 2 in Bueno de Mesquita *et al.* (2003, 144)'s Table 4.1).

Table 3: Savings, $n = 3942$

	Residuals		Controls	
Variable	Coef.	s.e.	Coef.	s.e.
W	12.64	(0.93)	-2.58	(1.51)
S	-1.43	(0.75)	5.69	(0.76)
ln(Population)	0.70	(0.13)	0.70	(0.13)
Democracy residuals	-7.74	(1.04)		
ln(GDP) residuals	6.24	(0.19)		
Democracy			-7.74	(1.04)
ln(GDP)			6.24	(0.19)
Constant	0.34	(2.10)	-39.94	(2.53)

When we substitute the actual controls for the residuals used by the authors, we see the coefficient on W changes signs. Equation 1 tells us why. Consider first the effect of the democracy residuals. The estimated coefficient in the regression of democracy on W, $(\mathbf{w}'\mathbf{w})^{-1}\mathbf{w}'\mathbf{d}$, is 1.01. $\hat{\beta}_D$, reading off Table 3, is -7.74. The bias is therefore $(\mathbf{w}'\mathbf{w})^{-1}\mathbf{w}'\mathbf{d}\hat{\beta}_2 = 1.01 * -7.74 = -7.82$. Now consider the effect of the GDP residuals. The estimated coefficient in the regression of GDP on W is 3.69. The effect of GDP, reading from the table, is 6.24. The bias in this case is $3.69 * 6.24 = 23.04$. The total bias is therefore $-7.82 + 23.04 = 15.22$, and the effect on the estimated W coefficient is $12.64 - (-7.82 + 23.04) = -2.58$, which is precisely what we find in column three of Table 3.[7]

We have learned a number of things from this exercise. First, replacing the residuals with democracy has a large effect on the coefficient on W when the

[7]The difference in the constants is a function of the different $\hat{\boldsymbol{\beta}}$s in the two equations and the "demeaning" of democracy and ln(GDP) in the residual regression. That is, Δ constant $= \bar{\mathbf{X}}_C\hat{\boldsymbol{\beta}}_C - \bar{\mathbf{X}}_R\hat{\boldsymbol{\beta}}_R$, where R stands for the residual regression and C stands for the control regression. In $\bar{\mathbf{X}}_R$, the means of democracy and ln(GDP) are set to zero.

effect of democracy on the dependent variable is large and has a concomitantly small effect when the effect of democracy is small. (This is because the first part of the bias, $(\mathbf{w}'\mathbf{w})^{-1}\mathbf{w}'\mathbf{d}$, the coefficient of W on democracy, is approximately 1.) Consequently, if democracy does not have a large effect on the dependent variable, using the procedure will create only a small bias, which may not affect our inferences from hypothesis tests.[8] Second, when the coefficient on democracy is negative, the coefficient on W decreases, rather than increases. Third, when the specification includes more than one set of residuals, the effect on the W coefficient is a combination of biases, and in the case of OLS, the effects are additively separable. In the example we gave in the previous paragraph, the negative effect that including the democratic residuals has on the W coefficient is more than offset by the positive effect that including the log income residuals has on the W coefficient.

Bueno de Mesquita *et al.* say they use the residualization technique to "assess the independent impact of democracy on our dependent variables," where the residuals are "the portion or characteristics of a country's degree of democracy or autocracy not endogenous to W and S." By "endogenous," the authors apparently mean that democracy is a function of W and S and thus these right-hand side variables are interdependent. Residualization was also used by Soper (1976) under exactly the same circumstances. He was criticized for doing so by Swan (1978) and Kennedy (1982) and many, if not all, of the same criticisms apply here as well.

If it is correct that democracy is dependent on W and S, this does not imply that the results of a regression of a dependent variable on W, S, and democracy are biased and inconsistent. Bueno de Mesquita *et al.*, writing in terms of endogeneity, in essence set up a system of simultaneous equations,

$$\begin{aligned} \mathbf{y} &= \beta_0 + \mathbf{w}\beta_1 + \mathbf{d}\beta_2 + \mathbf{x}\beta_3 + \boldsymbol{\epsilon} \\ \mathbf{d} &= \gamma_0 + \mathbf{w}\gamma_1 + \boldsymbol{\nu}, \end{aligned}$$

where $\mathbf{x}$ stands in for any of the other explanatory variables.

[8] Consider, for instance, the first regression in column 2 of Bueno de Mesquita *et al.*'s Table 4.1. Here, democracy does not have a large effect on log income; the coefficient is 0.34. Therefore, when we replace the residuals with democracy, the coefficient on W decreases only slightly.

This system of equations is recursive, and there is an endogeneity problem (an explanatory variable correlated with the error term) *only* if the two error terms, $\boldsymbol{\epsilon}$ and $\boldsymbol{\nu}$, are correlated (Greene 2003, 383). The question, then, is whether the error terms are independent. If the error terms are correlated, then the model Bueno de Mesquita *et al.* ran,

$$\mathbf{y} = \delta_0 + \mathbf{w}\delta_1 + \hat{\boldsymbol{\nu}}\beta_2 + \mathbf{x}\beta_3 + \boldsymbol{\epsilon},$$

makes little sense as $\hat{\boldsymbol{\nu}}$ and $\boldsymbol{\epsilon}$ *would still be correlated*. Use of the residualization technique does nothing to solve the endogeneity problem, if it exists.

There is a further problem with assuming that the error terms are correlated. If they are, then the upper equation in the system is unidentified because the endogenous variable in the lower equation is a linear function of an exogenous variable in the upper equation (Greene 2003, 395 and Davidson and MacKinnon 1981, 643). We can see the effect of this specification by noting that the upper equation fails to meet the order condition for identification.[9]

On the other hand, if we assume that the errors are uncorrelated and the system is recursive, then the correct estimator is straightforward multiple regression. Under these conditions, equation-by-equation OLS is the full information maximum likelihood estimator (Greene 2003, 397). Thus, whatever assumption we make about the error structure, residualization is not the appropriate technique for dealing with endogeneity.

Alternatively, and perhaps more sensibly, we could think of Bueno de Mesquita *et al.*'s concerns about the interdependence of their explanatory variables in terms of potential multicollinearity, which was also a concern of Soper's [42]. In fact, the residualization technique has traditionally been thought of as a "fix" for multicollinearity (see Goldberger 1964).

Although the likelihood that multicollinearity is a problem in these data is low (the number of observations is large and W and *democracy* are correlated at only 0.8), it is incumbent upon us to investigate the possibility that multicollinearity may affect our reestimation in Section 4. The reviewer's appendix contains collinearity diagnostics for all of our reestimations. The

[9] Another way to think about the model Bueno de Mesquita *et al.* ran is to substitute the equation for $\mathbf{d}$ into the equation for $\mathbf{y}$ and realize that the resulting equation is the reduced form (Swan 1978, 82). Of course, we care about the structural coefficients, which they do not report, and not the reduced form coefficients.

results show unequivocally that multicollinearity is not a problem in these data, and the results of our reestimations cannot be explained by it. Whether we think of the interrelatedness of winning coalition size and democracy as endogeneity or multicollinearity, multiple regression is the correct way to proceed.

3.1 A Brief History of Residualization

Bueno de Mesquita and colleagues are not alone in their use of this estimator. As noted earlier, the technique can be traced back to Goldberger (1964, 196), who wrote that given the regression, $\mathbf{y} = \mathbf{X}_1\boldsymbol{\beta}_1 + \mathbf{X}_2\boldsymbol{\beta}_2 + \boldsymbol{\epsilon}$, "regressing $\mathbf{y}$ on $\mathbf{X}_2$ and $\mathbf{X}_1$ together is equivalent to regression $\mathbf{y}$ on $\mathbf{X}_2$ corrected for $\mathbf{X}_1$, in the sense that it gives the same coefficient vector $\mathbf{b}_2$ for the regressors in $\mathbf{X}_2$." What many users fail to note is this is not true of the coefficient vector for the regressors in $\mathbf{X}_1$.

In economics, the first use of Goldberger's technique seems to be Ridker and Henning (1967), who are interested in the effects of air pollution on residential property values in the St. Louis metropolitan area. The bridge to political science came through economic studies of the effects of ideology on congressional voting (Kau and Rubin 1979; Kalt and Zupan 1984).[10] Kau and Rubin (1979) use the residualization technique to combat multicollinearity explicitly. They write, "To the extent that ADA does measure economic influence, including it in equations where the economic variables are already present may result in multicollinearity" [370]. Thus, they "residualize" the ADA scores by regressing the scores on the economic variables and then replacing the ADA scores in their regression model with the residuals, *RADA*. The result, as we have shown, is that the estimated coefficient on *RADA* is unbiased, but the estimated coefficients on the economic variables suffer from omitted variable bias. Kau and Rubin (1979), unlike Bueno de Mesquita *et al.*, replace their variable of interest with the residuals, thus ensuring an unbiased estimate. Bueno de Mesquita *et al.*, however, replace a control variable, democracy, with its residuals, thus ensuring their coefficient of interest is biased.

[10]Soper (1976)'s use of the residualization technique as a "fix" for endogeneity and multicollinearity does not seem to have been noticed beyond the discussion in *The Journal of Economic Education.*

Kalt and Zupan (1984) use a variation of the residualization technique in their study. They include the fitted values from the auxiliary regression along with residuals (Jackson and Kingdon 1992, 820). While the estimated coefficient on the residuals remains unbiased, the estimated coefficients on the other included variables are substantially more biased. Jackson and Kingdon (1992) criticize these studies on methodological grounds, but not for their use of the residualization technique. Rather, they argue that using interest group scores is predicting votes with votes and cannot capture member preferences. While Jackson and Kingdon note that the residualization procedure does not solve this particular problem, they do not directly address the residualization technique itself [813-814].

Variants of Kau and Rubin (1979) and Kalt and Zupan (1984) have been used by, among others, Johannes and McAdams (1981), Carson and Oppenheimer (1984), Fort, Hallagan, Morong, and Stegner (1993), Haider-Markel (1999), Uslaner (1999), Bishin (2000), and Schonhardt-Bailey (2003). Putting aside Jackson and Kingdon's critique, if these authors were solely interested in the coefficient on the ADA variable, the residualization technique might be appropriate given that they residualize the variable of interest. Most of these authors, however, are interested in the results of their statistical models beyond the ADA scores. Bishin (2000), for instance, is interested in a measure that reflects the economic interests of constituents. The estimated coefficient of this variable, however, suffers from omitted variable bias and is too large (constituent interest and ADA scores are positively correlated).

In the section to follow, we reestimate many of the results from Bueno de Mesquita *et al.* (2003) and show just how large an impact the residualization technique can have.

4 Replication and Reestimation Results

The attractive feature of Bueno de Mesquita *et al.* (2003)'s model is the wide range of observable implications that it generates. The authors gather a prodigious amount of data and conduct a series of ambitious statistical tests. Space does not allow us to replicate all of these results here, so we focus on chapters 4 and 5, which are the two core chapters of the book. The authors, in these chapters, test the implications of their theory for how leaders

of different political systems pursue different economic policies that lead to different growth trajectories and provide different levels of public goods. We replicate and reestimate all of the results in these two chapters, and we briefly present them. Few of their significant results survive reestimation when the residuals are replaced with the original variables.

Table 4: Reestimation Results for Chapter 4

Change in W Coefficient	**Not Significant** [†]	**Significant**[‡]
No change or larger		% Invest
Smaller		LogIncome[4.1]
Smaller: by at least 50%	Kleptocracy[4.8]	
Smaller: by at least 75%		
Change of Sign	% Savings Consumption LogIncome[4.4] Econ. Growth Kleptocracy[4.7]	Expenditures

Change in W/S Coefficient	**Not Significant**	**Significant**
No change or larger		% Invest
Smaller		LogIncome[4.1]
Smaller: by at least 50%	Kleptocracy[4.8]	
Smaller: by at least 75%	Expenditures	% Savings Consumption
Change of Sign	LogIncome[4.4]	Kleptocracy[4.7]

[†]Superscripts refer to the particular table when there is a question.

[‡]We treat 0.05 as the dividing line between significant and not.

4.1 Chapter 4: Economic Growth

Table 4 summarizes the results of reestimating the models in chapter 4 of Bueno de Mesquita *et al.* (2003). The models are categorized in Table 4 by the effect of W on the dependent variable when the models are reestimated. For instance, when we reestimate the effect of W on savings, the resulting coefficient is no longer statistically significant and has, in fact, changed signs. Reading from Table 4, we see that of the 17 coefficients reported by Bueno

de Mesquita *et al.* in chapter 4, 11 are no longer significant (9 cases) or have changed signs and are significant (2 cases). Of the remaining 6 regressions, 2 have coefficients that are reduced to the point where their substantive significance must be called into question. Thus, only 4 of the 17 regressions in chapter 4 continue to support the selectorate model. In the remainder of this section, we take a closer look at what Bueno de Mesquita *et al.* consider to be their most significant findings.

The central results in chapter 4 are the regressions that attempt to explain economic growth, and Bueno de Mesquita *et al.*'s results initially appear to be quite encouraging (Table 5, columns 1 and 3). They argue that "The structure of these tests makes achieving significance a challenge for any of the independent variables," and that "W and W/S are significant in their independent impact" (Bueno de Mesquita *et al.* 2003, 152). When analyzing models using the log of GDP as the dependent variable (which they interpret as a growth rate because they include a lagged dependent variable on the right-hand side), the authors particularly emphasize the substantive significance of the estimated coefficients, which range between 2.4 and 2.9 percent. They conduct a thought experiment based upon these rates of change that involves two hypothetical societies that begin with per capita income of $2,000 and a baseline real growth rate of zero. They conclude:

> A century of institutional difference results in the large-coalition system enjoying a real per capita income of $23,627, compared to $2,000 in real per capita income for the small-coalition system. These differences are so large that one might almost say that institutional differences by themselves seem sufficient to account for the disparities in real per capita income between otherwise seemingly comparable societies like North and South Korea, the former East and West Germany, or present Southern California and Northern Mexico (Bueno de Mesquita *et al.* 2003, 153).

Our reanalysis of the data does not support this conclusion. The reestimations are in Table 5 (columns 2 and 4).[11] The positive effects found by Bueno de Mesquita *et al.* are attributable to the bias created by the inclusion

[11] We present results only for the key variables W and W/S for space considerations and in keeping with the presentation of results in Bueno de Mesquita *et al.* (2003).

of the residuals, and these effects disappear when we control for the actual variables.

Table 5: Economic Growth, $n = 2931$

	ln(GDP per capita)							
	Residuals		Controls		Residuals		Controls	
Variable	Coef.	s.e.	Coef.	s.e.	Coef.	s.e.	Coef.	s.e.
W	0.029	(0.006)	-0.002	(0.008)				
W/S					0.024	(0.005)	-0.008	(0.007)

Note: All models include ln(Population) and either controls or residuals for civil liberties, political rights, investment, and savings.

Changing dependent variables does not save the selectorate model. In Table 6, we replicate and reestimate a model using economic growth rates reported by the World Bank as the dependent variable. Bueno de Mesquita *et al.* argue that political instability is costly (a negative coefficient for the variable *political change*), but widening the winning coalition offsets this (a positive coefficient for *change in* W/S) and having a wide coalition promotes growth (a positive coefficient for W).[12] The authors find a comforting policy recommendation in their results:

> At least from the perspective of the economic welfare of the average resident, the policy prescription is clear. In the short term, the enlargement of the winning coalition has on average either very mild negative consequences or positive benefits, and in the long run its growth advantages lift a country's residents out of poverty (Bueno de Mesquita *et al.* 2003, 157).

Unfortunately, Bueno de Mesquita *et al.*'s optimism about the long-term growth effects of W is not supported by the data. The steady-state effects of W on growth vanish when we substitute the actual variables for their residuals.[13] Our reestimation does confirm their analysis of the dynamics of regime change: democratization has no effect on growth rates, but moving

[12] *Political change* is "the squared difference in the coalition size W, compared with the coalition size W two years earlier"(Bueno de Mesquita *et al.* 2003, 154).

[13] In this particular regression, we have to substitute for the residuals on civil liberties, political rights, investment, and savings.

Table 6: GDP Growth, $n = 2312$

	Residuals		Controls	
Variable	Coef.	s.e.	Coef.	s.e.
W	1.73	(0.62)	-0.50	(0.82)
Political Change	-1.25	(0.45)	-1.25	(0.45)
Δ W/S	1.82	(0.76)	1.82	(0.76)

Note: All models include ln(Population) and either controls or residuals for civil liberties, political rights, investment, and savings.

from a democratic to an undemocratic regime has an effect that is significant and negative. This result is not a test of the selectorate theory, which makes no predictions about the effects of regime change.

On the basis of their results, Bueno de Mesquita *et al.* argue that it is significant that W, and not democracy, is responsible for promoting economic growth, and they draw an important policy conclusion from the distinction between W and democracy:

> If we add either the residual or the total effect of *Democracy* and *Democracy* squared to the regression results reported in table 4.5 [Table 6, first column in this paper], we do not find the significant downturn reported by Barro and others. More important, we do find that the effect of coalition size on the growth rate remains robust with regard to significance and the size of the coefficient. Thus, the effect of coalition size is robust and independent of other characteristics of democracy that may themselves dampen growth or have no consequential impact on growth. Institutional engineers in the World Bank, the IMF, and elsewhere would do well to consider this distinction when helping poorer countries reform their political institutions (Bueno de Mesquita, Smith, Siverson, and Morrow 2003, 160).

The data simply do not support this conclusion. Including democracy, with or without squared democracy, in the reestimation reported above (with actual variables rather than residuals) does not change the results qualita-

tively.[14] In the final analysis, W and W/S have little or no effect on economic growth, and the selectorate model receives little or no confirmation from the data.

4.2 Chapter 5: Public Goods, Social Services, and Private Benefits

In chapter 5, Bueno de Mesquita *et al.* (2003) test the implications of their model for the provision of "core public goods" (civil liberties, political rights, transparency, and peace), "general public goods" (public health, education, social security, and international trade), and finally private goods (black markets, construction, and corruption). The results of our chapter 5 reestimations are summarized in Table 7. In regressions that include residuals, Bueno de Mesquita *et al.* report significant estimates for the coefficients of W or W/S that support their hypotheses 53 times. Our reanalysis demonstrates that of these 53 coefficients, 22 are no longer significant or have the wrong sign when we substitute the original variables for the residuals. Of the 31 remaining coefficients, 18 (nearly 60%) are reduced in size by at least 50%, and seven are reduced in size by least 75%. In many of these cases, the reductions are sizeable enough to call into question the substantive significance of the variables. Thus, most of the evidence that Bueno de Mesquita *et al.* point to in support of their theory does not exist or is quite weak.

As in the previous section, we take a closer look at the results the authors point to as most important for their theory, and it is these key results that fare the worst. This is unsurprising as the bias produced by the residualization technique is most pronounced when the control variables are strongly related to the dependent variable. Thus, it is the dependent variables that are most strongly related to democracy and income for which the technique produces the most unreliable results. These variables tend to be the variables of greatest theoretical interest. We begin our discussion with core public goods.

The most important tests of their theory, Bueno de Mesquita *et al.* write, have to do with "core public goods":

[14]The estimated coefficient of W is -0.52 when polity is included in the regression reported in Table 4, column six, and -0.43 when polity and polity squared are included. Neither result is statistically significant ($p < 0.6$).

Table 7: Reestimation Results for Chapter 5

Change in W Coefficient	**Not Significant** [†]	**Significant**[†]
No change or larger		Doctors, Trade Measles Immunity
Smaller	Black Market Construction	Income[5], Low Birth Wt. Female Secondary Ed. DPT Immunity
Smaller: by at least 50%	Beds	Political Rights[5], Tax[2,5] Income[2], Death Rate Education Spending
Smaller: by at least 75%	Health Spending Corruption	Political Rights[2] Illiteracy Infant Mortality Life Expectancy
Smaller: by at least 95%	Civil Liberties[2,5] Yrs. of Education Water	
Change of Sign	War[2]	War[5], Social Security

Change in W/S Coefficient	**Not Significant**	**Significant**
No change or larger		Trade
Smaller	Black Market Construction	Female Secondary Ed. Low Birth Wt., Doctors Measles Immunity DPT Immunity
Smaller: by at least 50%		Tax, Income, Illiteracy Education Spending Death Rate
Smaller: by at least 75%	Health Spending Yrs. of Education Beds, Corruption	Political Rights Life Expectancy Infant Mortality
Change of Sign	Civil Liberties War, Water	Social Security

[†]Superscripts refer to the particular model.

> Civil liberties, political rights, transparency, peace and prosperity are among the most important public welfare enhancements that any government can provide.... Whatever else is found in the basket of public goods provided by government, the benefits of civil liberties, political rights, transparency, peace, and growing prosperity seem to be of universal desirability among residents of a state. Therefore, we treat civil liberties, political rights, transparency, and peace as core public goods (Bueno de Mesquita *et al.* 2003, 179).

The authors report results that appear to show a strong relationship between W and W/S and each of these key indicators of good governance: civil liberties, political rights, transparency, and peace. As Table 7 reveals, however, most of these results disappear in our reestimations, and all of the effects are substantially smaller than Bueno de Mesquita *et al.* believe.

The first set of variables examined are the Freedom House measures of civil liberties and political rights. These are often used as measures of democratic and authoritarian regime types, but, because they capture the behavioral dimensions of how regimes treat their citizens rather than the institutional differences among them, the authors regard them as measures of good government. Bueno de Mesquita *et al.* (2003, 181) conclude that "Both civil liberties and political rights show extraordinary dependence on coalition size." In our reestimations, however, the effects of W and W/S on civil liberties become insignificant, and their effects on political rights are dramatically reduced, diminishing their substantive impact. For example, the authors interpret their findings in this way:

> Eighty to ninety percent of the variance in these two fundamental benefits of a civil society is explained by institutional arrangements. Income also contributes strongly to the degree to which a society is characterized by civil liberties and political rights, but neither income effects nor the effects of democracy notably diminish the impact that coalition size has on these public goods...A shift from a score of 0 to a score of 1 on the indicator of coalition size, W, implies a four- to six-point jump on the seven-point civil liberties and political-rights scales (Bueno de Mesquita *et al.* 2003, 182).

Our reestimation shows, however, that a shift from the top to the bottom of the W scale leads to a shift of only one point on the seven-point political rights scale. The results are in Table 8. Thus, although the result remains statistically significant, W does not have a large substantive impact on political rights. Our finding is important because if Bueno de Mesquita *et al.* had been correct, it would have meant that the variation in W accounts for virtually the full range of variation in political rights. Such a finding would have been revolutionary. Though scholars of democracy debate the relative merits of institutional and behavioral definitions, they generally assume that these concepts are theoretically and empirically quite distinct. Bueno de Mesquita *et al.*'s finding, however, is an artifact of the biased estimation procedure, and the evidence simply does not support their conclusion.

Table 8: Political Rights, $n = 3223$

	Residuals		Controls		Residuals		Controls	
Variable	Coef.	s.e.	Coef.	s.e.	Coef.	s.e.	Coef.	s.e.
W	5.85	(0.057)	1.23	(0.090)				
W/S					5.87	(0.059)	1.20	(0.093)

Note: These models include ln(Population) and either controls or residuals for ln(GDP) and democracy.

Bueno de Mesquita *et al.* bolster their substantive claims by noting that for political rights, "the t-statistic varies between a low of 72 and a high of 102!" [182]. It is important to remember, however, that their estimation procedure, besides biasing the coefficient, also artificially decreases the standard errors. As we see in Table 8, the standard errors increase as the coefficients decrease. The t-statistics in the reestimations above are considerably more modest.

A similar story can be told of the effect of W and W/S on Bueno de Mesquita *et al.*'s measures of transparency, which are dummy variables indicating whether data on taxation and national income are reported to the World Bank. The biased estimation procedure once again leads the authors to exaggerate substantially the size of the effects. Writing of tax transparency, they claim that "The effects of coalition size are not only stronger than those for other political institutions, but they are also stronger than the independent impact of income on transparency" [183]. Similarly, writing of income reporting, they argue, "the effect of coalition size is substantially larger than the effect of other aspects of democratic or authoritarian governance" [184].

They draw these false conclusions because their estimates of these coefficients are more than twice as large as they should be.

Finally, we turn to war and peace, which are, of course, of primary interest to the authors. They ask whether countries with large coalitions and weak loyalty norms are less likely to go to war, and they find that W and W/S are strongly associated with peace:

> Whether alone, in concert with *WS:DemRes*, or in conjunction with *WS:DemRes* and *Parl_Pres*, large coalition size and a weak loyalty norm significantly reduce the probability of War (Bueno de Mesquita *et al.* 2003, 185).[15]

Table 9: War, $n = 1613$

	Residuals		Controls		Residuals		Controls	
Variable	Coef.	s.e.	Coef.	s.e.	Coef.	s.e.	Coef.	s.e.
W^2	-2.11	(0.649)	0.84	(0.546)				
W/S					-2.59	(0.703)	0.67	(0.536)
W^5	-1.61	(0.840)	1.84	(0.746)				

Note: These models include ln(Population) and either controls or residuals for ln(GDP) and democracy. Model 5 also contain *Parl_Pres*. Superscripts refer to the particular model.

These results do not survive our reestimation. The results are in Table 9. The effects of W and W/S reverse signs when we control for democracy. In two specifications they are insignificant, and in the third, the effect is significant in the wrong direction. On balance, then, the evidence suggests that large coalition size and a weak loyalty norm are likely to increase the probability war, not decrease it.

The biased substantive effects, almost as much as the incorrect hypothesis tests, lead Bueno de Mesquita *et al.* to draw erroneous inferences about the relationship between coalition size and democracy. In summing up their conclusions about the core public goods that they regard as the most important

[15] *WS:DemRes* are the residuals from a regression of democracy on W and S, and *Parl_Pres* are the residuals from a logit regression of W and S on a dummy variable for parliamentary regimes.

tests of their theory, Bueno de Mesquita *et al.* emphasize the substantive impact that coalition size appears to have and explicitly compare it to the substantive effect of democracy:

> The aspects of democracy or autocracy not accounted for by coalition size and selectorate size—except for one test regarding war—also significantly enhance the prospects that a government performs well in producing core public goods. However, in every instance, the strength of the association between coalition size and the provision of core public goods is stronger than the strength of association tied to other characteristics of a regime's institutions (Bueno de Mesquita *et al.* 2003, 186).

It is not far from this claim to the conclusion that coalition size is really the heart of the matter. Democracy, in short, is not as complex a matter as comparativists generally claim: its effects are largely reducible to a single parameter that is conceptually very straightforward. This inference, however, is based upon biased estimates that exaggerate the importance of coalition size by borrowing covariance from democracy.

Bueno de Mesquita *et al.* next consider the effects of W and W/S on a variety of public goods that they describe as "general public goods," such as public health, education, social security, and international trade. They emphasize that they do not regard any of these measures as sharp tests of their theory, but in aggregate they find that the results support their case:

> We believe that most people share a common desire to be governed by a regime that is attentive to providing the core public goods we have discussed. Other goods that potentially enhance the common welfare are more likely to be subject to variation in tastes and needs as we move from society to society. Some will prefer to emphasize education, others health care, and still others social security and so forth (Bueno de Mesquita, Smith, Siverson, and Morrow 2003, 186).

It is these general public goods, however, that most consistently support the hypotheses of the model. Three of four education measures and seven out

of ten public health measures remain statistically significant in our reestimations. Almost all of these effects, however, are substantially attenuated when we substitute actual variables for the residuals used by the authors. Their claims of strong effects, therefore, are generally incorrect. For example, the authors claim, "A shift from the smallest to the largest coalition structure produces a 20 to 40 percent reduction in the illiteracy rate among those fifteen years or older" [189]. These estimates are approximately three times larger than they should be (see Table 10). Similarly, the authors conclude, "living in a polity that depends on the largest-size coalition contributes nearly 13.5 more years of life [expectancy] than does living in a society that relies on the smallest-size coalitions" (Bueno de Mesquita *et al.* 2003, 194). These estimates are more than four times too large (see Table 10). These inflated substantive effects lead the authors to infer incorrectly, for example, that coalition size is the "key factor" that explains variations in public health [194].

Table 10: Public Goods, $n = 2709, 2042$

	Illiteracy				Life Expectancy			
	Residuals		Controls		Residuals		Controls	
Variable	Coef.	s.e.	Coef.	s.e.	Coef.	s.e.	Coef.	s.e.
W	-36.41	(1.30)	-12.55	(1.72)	13.45	(0.41)	2.74	(0.66)
W/S	-37.74	(1.30)	-10.10	(1.75)	13.67	(0.41)	2.72	(0.67)

Note: These models include ln(Population) and either controls or residuals for ln(GDP) and democracy.

Finally, Bueno de Mesquita *et al.* turn their attention to measures of private goods. The selectorate theory predicts that regimes with narrow coalitions and strong loyalty norms spend larger shares of national income on private benefits for the dictator and his or her close associates—indeed, the opportunity to spend resources to secure private benefits is why they spend less on public goods. The authors focus on three key measures of these private benefits: black markets, construction, and corruption. They argue that each of these represents opportunities for private enrichment: black markets exist because of government intervention or financial repression and create opportunities for arbitrage for the well-connected; construction is a convenient way to hide payoffs to friends and allies; corruption is a means of converting political influence into liquid assets. Their results demonstrate that each variable

responds in the predicted way to the size of coalitions and has dramatic substantive effects. None of these results survive reestimation, however; the results are in Table 11.

Table 11: Private Benefits, $n = 1254, 407, 1471$

	Residuals		Controls		Residuals		Controls	
Variable	Coef.	s.e.	Coef.	s.e.	Coef.	s.e.	Coef.	s.e.
BlackMarket								
W	-3.92	(1.74)	-3.13	(2.65)				
W/S					-3.88	(1.76)	-2.90	(2.64)
Corruption								
W	4.87	(0.39)	0.68	(0.69)				
W/S					4.92	(0.39)	0.73	(0.69)
Construction								
W	-12.43	(3.54)	-9.17	(6.27)				
W/S					-12.48	(3.57)	-9.24	(6.16)

Note: These models include ln(Population) and either controls or residuals for ln(GDP) and democracy.

In the case of corruption, the substantive significance of the estimated coefficient drops almost to zero. In the other cases, while the estimated coefficients shrink, the change is not large. There is a substantial change, however, in the standard errors. For black market, the standard errors are 50% larger in the reestimation. For construction, they are over 70% larger. As a result, rather than support Bueno de Mesquita *et al.*'s theory, these results support the null hypothesis of no effect. There is little statistical evidence to suggest that W or W/S has an impact on any of these measures of private goods.

In summary, Bueno de Mesquita *et al.* report 70 regressions that use the procedure that we criticize. When we reestimate these regressions and substitute the original control variables for their residuals, we find that in 33 of those 70 analyses, or 47% of the time, the biased estimator incorrectly confirmed a false hypothesis. Of the remaining 37 analyses, the biased procedure exaggerates the substantive effects of W and W/S by at least a factor of 2 in 20 of the regressions, or 54% of the time. In 53 of 70 regressions, then, the evidence that Bueno de Mesquita *et al.* marshal in support of their theory does not exist.

There are a few cases (investment, trade, the number of doctors, and im-

munity to measles) where the biased procedure works against the authors and makes their results appear to be weaker than they are. However, the results that the authors indicate are the most important are all overturned. These key findings include the results on government expenditures and economic growth in chapter 4 and the results on core public goods in chapter 5. Bueno de Mesquita *et al.* obtained their misleading results because W is correlated with democracy, and the biased procedure they use exploits this fact to inflate the apparent effect of W and decrease its standard error. Even in cases where W remains statistically significant, its actual effects turn out to be quite modest. The combination of W's apparent invincibility—it seemed to never fail a hypothesis test—and its inflated substantive effects led the authors to conclude that coalition size is the core of the matter where democracy is concerned. What appeared to be a conceptually cleaner measure that performed better than democracy was really the artifact of a biased estimation technique.

5 A Better Measure

We noted above that one possible interpretation of the findings in the last section is that the data collected by Bueno de Mesquita *et al.* are inadequate to test the theory. Since their measures of W and S do not directly measure the phenomena of coalition size or selectorate size, it is possible both for the theory to be true and for W and S to have no significant effects. To test this conjecture, we reestimate Bueno de Mesquita *et al.*'s regressions using a theoretically more appropriate measure of one of their key variables, the ratio of coalition size to selectorate size, W/S.

That the data used in *The Logic of Political Survival* are inadequate to test the theory is plausible. The measure of coalition size, W, is constructed from four dummy variables: three components of the polity score, *competitivesness of executive recruitment*, *openness of executive recruitment*, and *competitiveness of participation*, plus one derived from a variable from Arthur Banks's data, *civilian character of regime*. These four variables are summed, the result is normalized to (0, 1), and the resulting variable takes five possible values. A maximum value indicates that a government is not a military regime, that the executive is not selected in unopposed elections and is not hereditary, and that there are stable political groups that compete for influence.

In short, a country achieves a 1 on the W scale if it satisfies four necessary conditions to be considered a democracy, but this tells us little about the size of the winning coalition required by its political institutions. The measure chosen for S is a normalization to (0, 1) of the *legislative selection* variable from Polity II, which takes three values depending upon whether a legislature exists, and if so, whether it is appointed or elected. Again, all democracies should score a 1 (Bueno de Mesquita *et al.* 2003, 134-135).

Table 12: Results when W/S is replaced by Votegov.

Table	**Supports**	**Opposes**	**Insignificant**	**Insufficient**
Table 4.1	LogIncome		% Invest % Savings	Consumption
Table 4.4			LogIncome	
Table 4.6		Expenditures		
Table 4.7		Kleptocracy		
Table 4.8				Kleptocracy
Table 5.1			Civil Liberties Political Rights	
Table 5.2			Tax	Income
Table 5.3		War		
Table 5.5		Education Spending Female Secondary Ed.	Yrs. of Education Illiteracy	
Table 5.6		Infant Mortality Measles Immunity DPT Immunity Social Security	Life Expectancy Death Rate Health Spending Doctors Beds Low Birth Wt. Water	
Table 5.7	Trade			
Table 5.8		Black Market	Corruption Construction	

These measures are not sensitive enough to discriminate among the variety of political institutions in the way Bueno de Mesquita *et al.* desire. While it is true that each component of the W score is associated with wider winning coalitions, this is because democracies have wider winning coalitions

than autocracies. W achieves more variation than a dichotomous measure of democracy only because it discriminates roughly among alternative kinds of autocratic governments. It is not clear, however, that these variations correspond to increments in winning coalition size. Furthermore, these variables do not allow the authors to test the claims they make about discriminating the effect of W from that of W/S, because the correlation between these two variables is 0.99. For statistical purposes, their measures of W and W/S are identical.

Fortunately, an alternative measure of W/S exists, which is ideally suited to test Bueno de Mesquita *et al.*'s hypotheses, although it was gathered with other purposes in mind. In his study of representation in democracies, Powell (2000) is motivated by a theoretical interest in determining how alternative electoral institutions affect the degree to which elections represent the preferences of the electorate in the political process. In particular, he argues that a trade-off exists between the ability of single-member district (SMD) electoral systems to generate decisive outcomes and stable majority governments and the ability of proportional representation (PR) systems to broadly represent the preferences of the electorate. In order to measure the degree of representation, therefore, he constructs a variable that measures the proportion of the electorate that voted for parties that subsequently joined the governing coalition. This is a direct measure of coalition size divided by electorate size, or W/S. The data exist for twenty countries and elections ranging from 1967 to 1994.

In the results reported in Table 12, we use the Powell measure of representation as a measure of W/S. Of the 32 cases in which we are able to replicate regressions using W/S, only two provide evidence that supports Bueno de Mesquita *et al.*'s hypotheses. On the other hand, ten provide statistically significant results that reject their hypotheses, and seventeen provide insignificant results. In three cases there were insufficient data to represent a meaningful test.

6 Conclusion

Bueno de Mesquita and coauthors have made an important contribution by providing a theoretical model that generates numerous testable hypotheses

and then subjecting them to a wide range of empirical tests. Furthermore, they have followed best practices by making their data publicly available, without which assistance this reanalysis of their results would have been practically impossible.

Unfortunately, our reanalysis indicates that Bueno de Mesquita *et al.*'s enthusiasm about the predictive power of their key variables, W and W/S, is unwarranted. The size of a country's selectorate and winning coalition are not associated with the decisions made by leaders. The authors are misled because all of their statistical evidence suffers from induced omitted variable bias. The authors do not control for democracy and income, but rather, they control for the residuals of democracy and income from auxiliary regressions of these variables on W or W/S. The effect of this residualization on their coefficients of interest is precisely that of omitted variable bias. Their estimates of W and W/S are therefore both biased and inconsistent.

The correct way to run these regressions is to substitute the original variables for the residuals. When we do that, the majority of Bueno de Mesquita *et al.*'s findings disappear. Thus, the main effects that they attribute to W and W/S are due to their correlation with other variables, and the apparent robustness of their results to the inclusion of a variety of control variables is attributable to the biased estimation procedure. Our results cannot be attributed to multicollinearity as there is no evidence of multicollinearity in any of our regressions. It is the estimation procedure that is pernicious, and we hope the present study convinces others to avoid using it in the future.

Up to this point, we were unable to distinguish between two possibilities: that the theory is false; and that the available data are inadequate to test the theory. The authors acknowledge that their key variables are crude and do not directly measure the quantities of interest, but at the same time, claim that their data are nevertheless adequate to test their claims. As a check, we repeated as many tests as possible using an alternative measure of coalition size as a proportion of the electorate carefully developed by Powell (2000). Again, we found virtually no evidence to support Bueno de Mesquita *et al.*'s hypotheses and substantial refuting evidence.

The Logic of Political Survival makes the arresting claim that it has isolated the key mechanism by which democracy generates its benefits, thereby resolving the debate between the advocates of institutions, behavior, and political culture. The empirical evidence, however, does not support this claim.

Democracy remains a complex and rather inscrutable phenomenon that is subject to multiple interpretations. Future scholars will no doubt find Bueno de Mesquita *et al.*'s model a useful starting point, but they must build on it with the understanding that there is little empirical evidence to support it.

A Proofs

In what follows, we abstract from Bueno de Mesquita *et al.*'s model slightly by ignoring the panel data set-up and the additional control variables. Our statistical conclusions are unaffected by these simplifications.

A.1 Expectation and Consistency

The true equation is

$$\mathbf{y} = \beta_0 + \mathbf{w}\beta_1 + \mathbf{d}\beta_2 + \boldsymbol{\epsilon},$$

where $\mathbf{y}$ is an $n \times 1$ vector of observations on any number of dependent variables, $\mathbf{w}$ is an $n \times 1$ vector of observations on *winning coalition size*, and $\mathbf{d}$ is an $n \times 1$ vector of observations on *level of democracy*, and $\boldsymbol{\epsilon}$ is an $n \times 1$ vector of disturbances. Concerned about multicollinearity, Bueno de Mesquita *et al.* substitute the residuals, $\hat{\boldsymbol{\nu}}$, from a regression of *level of democracy* on *winning coalition size* ($\mathbf{d} = \mathbf{w}\gamma + \boldsymbol{\nu}$) for *level of democracy*,

$$\begin{aligned} \mathbf{y} &= \beta_0 + \mathbf{w}\beta_1 + \hat{\boldsymbol{\nu}}\beta_2 + \boldsymbol{\epsilon} \\ &= \beta_0 + \mathbf{w}\beta_1 + \mathbf{M}\mathbf{d}\beta_2 + \boldsymbol{\epsilon}, \end{aligned}$$

where $\mathbf{M} = (\mathbf{I} - \mathbf{w}(\mathbf{w}'\mathbf{w})^{-1}\mathbf{w}')$. The estimated coefficients are

$$\begin{aligned} \begin{pmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \end{pmatrix} &= \begin{pmatrix} \mathbf{w}'\mathbf{w} & \mathbf{w}'\mathbf{M}\mathbf{d} \\ \mathbf{d}'\mathbf{M}\mathbf{w} & \mathbf{d}'\mathbf{M}\mathbf{d} \end{pmatrix}^{-1} \begin{pmatrix} \mathbf{w}' \\ \mathbf{d}'\mathbf{M} \end{pmatrix} [\mathbf{w}\beta_1 + \mathbf{d}\beta_2 + \boldsymbol{\epsilon}] \\ &= \begin{pmatrix} \mathbf{w}'\mathbf{w} & 0 \\ 0 & \mathbf{d}'\mathbf{M}\mathbf{d} \end{pmatrix}^{-1} \begin{pmatrix} \mathbf{w}' \\ \mathbf{d}'\mathbf{M} \end{pmatrix} [\mathbf{w}\beta_1 + \mathbf{d}\beta_2 + \boldsymbol{\epsilon}] \\ &= \begin{bmatrix} (\mathbf{w}'\mathbf{w})^{-1} & 0 \\ 0 & (\mathbf{d}'\mathbf{M}\mathbf{d})^{-1} \end{bmatrix} \begin{pmatrix} \mathbf{w}'\mathbf{w}\beta_1 + \mathbf{w}'\mathbf{d}\beta_2 + \mathbf{w}'\boldsymbol{\epsilon} \\ \mathbf{d}'\mathbf{M}\mathbf{w}\beta_1 + \mathbf{d}'\mathbf{M}\mathbf{d}\beta_2 + \mathbf{d}'\mathbf{M}\boldsymbol{\epsilon} \end{pmatrix}. \end{aligned}$$

Taking expectations, with $E[\boldsymbol{\epsilon}] = \mathbf{0}$,

$$E\begin{pmatrix}\hat{\beta}_1\\ \hat{\beta}_2\end{pmatrix} = \begin{pmatrix}\beta_1\\ \beta_2\end{pmatrix} + \begin{pmatrix}(\mathbf{w}'\mathbf{w})^{-1}\mathbf{w}'\mathbf{d}\beta_2\\ 0\end{pmatrix}$$

because $\mathbf{d}'\mathbf{M}\mathbf{w} = 0$.

The inconsistency of $\hat{\beta}_1$ follows from the fact that the bias on $\hat{\beta}_1$ is precisely that of omitted variable bias.

A.2 Variance

The estimated coefficient are, from above,

$$\begin{pmatrix}\hat{\beta}_1\\ \hat{\beta}_2\end{pmatrix} = \begin{pmatrix}\beta_1\\ \beta_2\end{pmatrix} + \begin{pmatrix}(\mathbf{w}'\mathbf{w})^{-1}\mathbf{w}'\mathbf{d}\beta_2 + (\mathbf{w}'\mathbf{w})^{-1}\mathbf{w}'\boldsymbol{\epsilon}\\ (\mathbf{d}'\mathbf{M}\mathbf{d})^{-1}\mathbf{d}'\mathbf{M}\boldsymbol{\epsilon}\end{pmatrix}.$$

Thus, we can write the variance as

$$\begin{aligned} E\begin{bmatrix}\hat{\beta}_1 - E(\hat{\beta}_1)\\ \hat{\beta}_2 - E(\hat{\beta}_2)\end{bmatrix}\begin{bmatrix}\hat{\beta}_1 - E(\hat{\beta}_1)\\ \hat{\beta}_2 - E(\hat{\beta}_2)\end{bmatrix}' &= \begin{pmatrix}(\mathbf{w}'\mathbf{w})^{-1}\mathbf{w}'E[\boldsymbol{\epsilon}\boldsymbol{\epsilon}']\mathbf{w}(\mathbf{w}'\mathbf{w})^{-1}\\ (\mathbf{d}'\mathbf{M}\mathbf{d})^{-1}\mathbf{d}'\mathbf{M}E[\boldsymbol{\epsilon}\boldsymbol{\epsilon}']\mathbf{M}\mathbf{d}(\mathbf{d}'\mathbf{M}\mathbf{d})^{-1}\end{pmatrix}\\ \mathrm{Var}\begin{pmatrix}\hat{\beta}_1\\ \hat{\beta}_2\end{pmatrix} &= \begin{pmatrix}\sigma^2(\mathbf{w}'\mathbf{w})^{-1}\\ \sigma^2(\mathbf{d}'\mathbf{M}\mathbf{d})^{-1}\end{pmatrix}. \end{aligned}$$

We can compare the variance of the estimated coefficients from the Bueno de Mesquita *et al.* model to the variance of the estimated coefficients of the correctly specified model by comparing their inverses (Greene 2003, 150). Let the estimated coefficients from the correctly specified model be $\hat{\beta}_{i.c}$. Then,

$$\begin{aligned} \mathrm{Var}\begin{pmatrix}\hat{\beta}_1\\ \hat{\beta}_2\end{pmatrix}^{-1} - \mathrm{Var}\begin{pmatrix}\hat{\beta}_{1.c}\\ \hat{\beta}_{2.c}\end{pmatrix}^{-1} &= \frac{1}{\sigma^2}\begin{pmatrix}\mathbf{w}'\mathbf{w} - [\mathbf{w}'\mathbf{w} - \mathbf{w}'\mathbf{M}\mathbf{d}(\mathbf{d}'\mathbf{M}\mathbf{d})^{-1}\mathbf{d}'\mathbf{M}\mathbf{w}]\\ (\mathbf{d}'\mathbf{M}\mathbf{d}) - (\mathbf{d}'\mathbf{M}\mathbf{d})\end{pmatrix}\\ &= \frac{1}{\sigma^2}\begin{pmatrix}\mathbf{w}'\mathbf{M}\mathbf{d}(\mathbf{d}'\mathbf{M}\mathbf{d})^{-1}\mathbf{d}'\mathbf{M}\mathbf{w}\\ 0\end{pmatrix}. \end{aligned}$$

As $(1/\sigma^2)\mathbf{w}'\mathbf{M}\mathbf{d}(\mathbf{d}'\mathbf{M}\mathbf{d})^{-1}\mathbf{d}'\mathbf{M}\mathbf{w}$ is nonnegative definite, the variance of $\hat{\beta}_1$ is never larger than the variance of $\hat{\beta}_{1.c}$. Thus, the estimator is biased, and the variance is too small. The variance of $\hat{\beta}_2$ remains unaffected.

A.3 Mean Squared Error Comparison

The mean square error for the correct estimator is[16]

$$\begin{aligned}
\mathrm{MSE}_{\hat{\beta}_{1.c}} &= \sigma^2(\mathbf{w}'\mathbf{M}_d\mathbf{w})^{-1} \\
&= \sigma^2(\mathbf{w}'\mathbf{w} - \mathbf{w}'\mathbf{d}(\mathbf{d}'\mathbf{d})^{-1}\mathbf{d}'\mathbf{w})^{-1} \\
&= \sigma^2\{(\mathbf{w}'\mathbf{w})^{-1} + (\mathbf{w}'\mathbf{w})^{-1}\mathbf{w}'\mathbf{d}[\mathbf{d}'\mathbf{d} - \mathbf{d}'\mathbf{w}(\mathbf{w}'\mathbf{w})^{-1}\mathbf{w}'\mathbf{d}]^{-1}\mathbf{d}'\mathbf{w}(\mathbf{w}'\mathbf{w})^{-1}\}.
\end{aligned}$$

where $\mathbf{M}_d = \mathbf{I} - \mathbf{d}(\mathbf{d}'\mathbf{d})^{-1}\mathbf{d}'$.

The mean square error for the residualization estimator is

$$\mathrm{MSE}_{\hat{\beta}_1^*} = \sigma^2(\mathbf{w}'\mathbf{w})^{-1} + (\mathbf{w}'\mathbf{w})^{-1}\mathbf{w}'\mathbf{d}\beta_2^2\mathbf{d}'\mathbf{w}(\mathbf{w}'\mathbf{w})^{-1}.$$

The difference between these measures is

$$\begin{aligned}
\Delta &= \mathrm{MSE}_{\hat{\beta}_1} - \mathrm{MSE}_{\hat{\beta}_1^*} \\
&= \sigma^2(\mathbf{w}'\mathbf{w})^{-1}\mathbf{w}'\mathbf{d}[\mathbf{d}'\mathbf{d} - \mathbf{d}'\mathbf{w}(\mathbf{w}'\mathbf{w})^{-1}\mathbf{w}'\mathbf{d}]^{-1}\mathbf{d}'\mathbf{w}(\mathbf{w}'\mathbf{w})^{-1} \\
&\quad -(\mathbf{w}'\mathbf{w})^{-1}\mathbf{w}'\mathbf{d}\beta_2^2\mathbf{d}'\mathbf{w}(\mathbf{w}'\mathbf{w})^{-1} \\
&= (\mathbf{w}'\mathbf{w})^{-1}\mathbf{w}'\mathbf{d}\{\sigma^2[\mathbf{d}'\mathbf{d} - \mathbf{d}'\mathbf{w}(\mathbf{w}'\mathbf{w})^{-1}\mathbf{w}'\mathbf{d}]^{-1} - \beta_2^2\}\mathbf{d}'\mathbf{w}(\mathbf{w}'\mathbf{w})^{-1} \\
&= (\mathbf{w}'\mathbf{w})^{-1}\mathbf{w}'\mathbf{d}[\sigma^2(\mathbf{d}'\mathbf{M}_w\mathbf{d})^{-1} - \beta_2^2]\hat{\mathbf{d}}'\mathbf{w}(\mathbf{w}'\mathbf{w})^{-1}
\end{aligned}$$

where $\mathbf{M}_w = \mathbf{I} - \mathbf{w}(\mathbf{w}'\mathbf{w})^{-1}\mathbf{w}'$. The first term inside the brackets is the covariance matrix for $\hat{\beta}_2$. When the full expression inside the brackets is positive semidefinite, then Δ is positive semidefinite, and the residualization estimator is to be preferred. To prefer the residualization estimator would mean that the variance of $\hat{\beta}_2$, $\sigma^2(\mathbf{d}'\mathbf{M}_w\mathbf{d})^{-1}$, is larger than β_2 squared. This situation is highly unlikely given that democracy is an important variable.

[16] Using the result that $[\mathbf{A} + \mathbf{BDC}]^{-1} = \mathbf{A}^{-1} - \mathbf{A}^{-1}\mathbf{B}(\mathbf{CA}^{-1}\mathbf{B} + \mathbf{D}^{-1})^{-1}\mathbf{CA}^{-1}$.

References

Almond, Gabriel A., and Sidney Verba. 1963. *The Civic Culture: Political Attitudes and Democracy in Five Nations*. Princeton: Princeton University Press.

Bishin, Benjamin G. 2000. "Constituency Influence in Congress: Does Subconstituency Matter?" *Legislative Studies Quarterly* 25:389–415.

Bueno de Mesquita, Bruce, Alastair Smith, Randolph M. Siverson, and James D. Morrow. 1999. "An Institutional Explanation of the Democratic Peace." *American Political Science Review* 93:791–807.

Bueno de Mesquita, Bruce, Alastair Smith, Randolph M. Siverson, and James D. Morrow. 2003. *The Logic of Political Survival*. Cambridge: The MIT Press.

Carson, Richard T., and Joe A. Oppenheimer. 1984. "A Method of Estimating the Personal Ideology of Political Representatives." *American Poltiical Science Review* 78:163–178.

Cox, Gary W. 1997. *Making Votes Count: Strategic Coordination in the Worlds Electoral Systems*. New York: Cambridge University Press.

Dahl, Robert A. 1971. *Polyarchy: Participation and Opposition*. New Haven: Yale University Press.

Davidson, Russell, and James G. MacKinnon. 1981. "Several Tests for Model Specification in the Presence of Alternative Hypotheses." *Econometrica* 49:781–793.

Fort, Rodney, William Hallagan, Cyril Morong, and Tesa Stegner. 1993. "The Ideological Component of Senate Voting: Difference Principles or Different Principals?" *Public Choice* 76:39–57.

Goldberger, Arthur S. 1964. *Econometric Theory*. New York: Wiley.

Greene, William H. 2003. *Econometric Analysis*. 5 ed. New Jersey: Prentice Hall.

Haider-Markel, Donald P. 1999. "Redistributing Values in Congress: Interest Group Influence Under Sub-Optimal Conditions." *Political Research Quarterly* 52:113–144.

Hicken, Allen, Shanker Satyanath, and Ernest Sergenti. 2005. "Political Institutions and Economic Performance: The Effects of Accountabil-

ity and Obstacles to Policy Change." *American Journal of Political Science* 49:897–907.

Jackson, John E., and John W. Kingdon. 1992. "Ideology, Interest Group Scores, and Legislative Votes." *American Journal of Political Science* 36:805–823.

Johannes, John R., and John C. McAdams. 1981. "The Congressional Incumbency Effect: Is it Casework, Policy Compatibility, or Something Else? An Examination of the 1978 Election." *American Political Science Review* 25:512–542.

Kalt, Joseph P., and Mark A. Zupan. 1984. "Capture and Ideology in the Economic Theory of Politics." *American Economic Review* 74:279–300.

Kau, James B., and Paul H. Rubin. 1979. "Self-Interest, Ideology, and Logrolling in Congressional Voting." *Journal of Law and Economics* 22:365–384.

Kennedy, Peter E. 1982. "Eliminating Problems Caused by Multicollinearity: A Warning." *The Journal of Economic Education* 13:62–64.

Lijphart, Arend. 1999. *Patterns of Democracy: Government Forms and Performance in Thirty-Six Countries*. New Haven: Yale University Press.

Powell, G. Bingham. 2000. *Elections as Instruments of Democracy*. New Haven: Yale University Press.

Przeworski, Adam, Michael E. Alvarez, Jose Antonio Cheibub, and Fernando Limongi. 2000. *Democracy and Development: Political Institutions and Well-Being in the World, 1950-1990*. New York: Cambridge University Press.

Putnam, Robert D. 1994. *Making Democracy Work: Civic Traditions in Modern Italy*. Princeton: Princeton University Press.

Ridker, Ronald G., and John A. Henning. 1967. "The Determinants of Residential Property Values with Special Reference to Air Pollution." *The Review of Economics and Statistics* 49:246–257.

Schonhardt-Bailey, Cheryl. 2003. "Ideology, Party and Interests in the British Parliament of 1841-47." *British Journal of Political Science*

33:581–605.

Shugart, Matthew S., and John M. Carey. 1992. *Presidents and Assemblies: Constitutional Design and Electoral Dynamics*. New York: Cambridge University Press.

Soper, John C. 1976. "Second Generation Research in Economic Education: Problems of Specification and Interdependence." *The Journal of Economic Education* 8:40–48.

Swan, Craig. 1978. "Comments on the Problems of Specification and Interdependence in a Set of Learning Equations." *The Journal of Economic Education* 9:81–86.

Uslaner, Eric M. 1999. *The Movers and the Shrikers: Representatives and Ideologues in the Senate*. Ann Arbor: The University of Michigan Press.